

## Fallacies in the interpretation of text-matching software similarity reports

*Salim Razi*

*Canakkale Onsekiz Mart University, Turkey*

**Type:** Workshop

**Section:** Use of technology in addressing issues of academic integrity

**Abstract:**

In the digital era, detection of plagiarism often depends on text-matching software similarity reports. Although such reports have the potential to facilitate the detection of plagiarism, they should be approached with caution. All text-matching software aims to reveal plagiarism concerns by highlighting matches between the submitted text and other sources included in its database. To contribute to the interpretation of similarity reports, it is also a common practice for text-matching software to attach a similarity ratio. Although the ratio is intended to provide supportive statistical information about the matches, in practice such statistics are often treated as the main concern, and it is quite a common practice for lecturers simply to depend on the statistics rather than check whether the highlighted expressions are due to plagiarism infringements or not. Within this respect, several hard-copy similarity report samples will be delivered to the audience in this workshop session to exemplify how isolated statistics from highlighted matches may mislead lecturers. These samples will illustrate why academic practices that merely depend on pure statistics follow a flawed methodology in the detection of plagiarism. The audience will be provided with an ideal similarity report interpretation procedure which is expected to contribute to the development of policies, especially in institutions where plagiarism is wrongly acceptable until a certain threshold similarity ratio.

**Keywords:** Detecting plagiarism, plagiarism policies, similarity reports, text-matching software.